



OIPF

## RAW SEQUENCE LISTING

DATE: 01/30/2002

PATENT APPLICATION: US/10/034,621

TIME: 18:26:31

Input Set : N:\Crf3\RULE60\10034621.raw

Output Set: N:\CRF3\01302002\J034621.raw

1 <110> APPLICANT: DIVERSA CORPORATION  
 2 CALLEN, Walter  
 3 MATHUR, Eric  
 4 <120> TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY  
 5 AND METHODS OF USE  
 6 THEREOF  
 7 <130> FILE REFERENCE: DIVER1350-2  
 8 <140> CURRENT APPLICATION NUMBER: 10/034,621  
 C--> 9 <141> CURRENT FILING DATE: 2000-12-21  
 11 <150> PRIOR APPLICATION NUMBER: US/09/656,309  
 12 <151> PRIOR FILING DATE: 2000-09-06  
 15 <160> NUMBER OF SEQ ID NOS: 2  
 16 <170> SOFTWARE: PatentIn version 3.0  
 18 <210> SEQ ID NO: 1  
 19 <211> LENGTH: 2412  
 20 <212> TYPE: DNA  
 21 <213> ORGANISM: Pyrolobus fumarius  
 22 <220> FEATURE:  
 23 <221> NAME/KEY: CDS  
 24 <222> LOCATION: (1)..(2412)  
 25 <400> SEQUENCE: 1  
 26 atg act gaa gtt gta ttc acg gtt tta gac tct agc tac gag gtt gtt 48  
 27 Met Thr Glu Val Val Phe Thr Val Leu Asp Ser Ser Tyr Glu Val Val  
 28 1 5 10 15  
 29 ggt aaa gag cct cag gta atc ata tgg ggt att gct gag aac ggc gag 96  
 30 Gly Lys Glu Pro Gln Val Ile Ile Trp Gly Ile Ala Glu Asn Gly Glu  
 31 20 25 30  
 32 agg gta gtc ctc att gac agg tct ttt cgc cca tac ttc tat gcg ctg 144  
 33 Arg Val Val Leu Ile Asp Arg Ser Phe Arg Pro Tyr Phe Tyr Ala Leu  
 34 35 40 45  
 35 ctt gca ccg ggc gcc gat cct aag cag gta gca caa cgt att cgt gca 192  
 36 Leu Ala Pro Gly Ala Asp Pro Lys Gln Val Ala Gln Arg Ile Arg Ala  
 37 50 55 60  
 38 ttg agt agg cca aag agc ccg att ata ggt gta gag gat gac aag agg 240  
 39 Leu Ser Arg Pro Lys Ser Pro Ile Ile Gly Val Glu Asp Asp Lys Arg  
 40 65 70 75 80  
 41 aag tac ttc ggg agg cct cgt agg gtc tta cgt att cgc acc gtg cta 288  
 42 Lys Tyr Phe Gly Arg Pro Arg Arg Val Leu Arg Ile Arg Thr Val Leu  
 43 85 90 95  
 44 ccc gag gct gtt agg gag tat cgc gaa ctc gta aag aac gtt gat ggt 336  
 45 Pro Glu Ala Val Arg Glu Tyr Arg Glu Leu Val Lys Asn Val Asp Gly  
 46 100 105 110  
 47 gtt gag gat gtt cta gag gcg gat ata cgc ttc gct atg cgc tat ctc 384

ENTERED

## RAW SEQUENCE LISTING

DATE: 01/30/2002

PATENT APPLICATION: US/10/034,621

TIME: 18:26:31

Input Set : N:\Crf3\RULE60\10034621.raw

Output Set: N:\CRF3\01302002\J034621.raw

48	Val Glu Asp Val Leu Glu Ala Asp Ile Arg Phe Ala Met Arg Tyr Leu	
49	115 120 125	
50	ata gat cac gat cta ttt cct ttc acc tgg tac cgt gta gag gct gag	432
51	Ile Asp His Asp Leu Phe Pro Phe Thr Trp Tyr Arg Val Glu Ala Glu	
52	130 135 140	
53	ccc ctc gag aac aag atg ggc ttc cgt gtc gac aag gta tac ctg gtt	480
54	Pro Leu Glu Asn Lys Met Gly Phe Arg Val Asp Lys Val Tyr Leu Val	
55	145 150 155 160	
56	aag agc agg ccg gag cca ctt tat ggt gag gct ctc gca cca acc aag	528
57	Lys Ser Arg Pro Glu Pro Leu Tyr Gly Glu Ala Leu Ala Pro Thr Lys	
58	165 170 175	
59	ctt ccc gat ctt agg ata ctc gcg ttc gat att gaa gtt tat agc aag	576
60	Leu Pro Asp Leu Arg Ile Leu Ala Phe Asp Ile Glu Val Tyr Ser Lys	
61	180 185 190	
62	caa ggg tcg ccg cgt cca gag cgc gat cct gta ata gtg ata gct gtg	624
63	Gln Gly Ser Pro Arg Pro Glu Arg Asp Pro Val Ile Val Ile Ala Val	
64	195 200 205	
65	aag act gac gat ggc gat gag gtg cta ttc att gca gag ggc aaa gac	672
66	Lys Thr Asp Asp Gly Asp Glu Val Leu Phe Ile Ala Glu Gly Lys Asp	
67	210 215 220	
68	gat cga aaa ccg ata cgc gag ttt gta gag tac gtg aag agg tat gac	720
69	Asp Arg Lys Pro Ile Arg Glu Phe Val Glu Tyr Val Lys Arg Tyr Asp	
70	225 230 235 240	
71	ccc gac ata ata gtc ggt tat aac aac aat cat ttc gat tgg cct tat	768
72	Pro Asp Ile Ile Val Gly Tyr Asn Asn Asn His Phe Asp Trp Pro Tyr	
73	245 250 255	
74	ctt ttg agg cgc gcc cgc atc cta ggc ata aag ctt gat gtg act aga	816
75	Leu Leu Arg Arg Ala Arg Ile Leu Gly Ile Lys Leu Asp Val Thr Arg	
76	260 265 270	
77	aga gtt ggc gcc gag ccc acc act agc gta cat ggg cac gtc tct gtc	864
78	Arg Val Gly Ala Glu Pro Thr Thr Ser Val His Gly His Val Ser Val	
79	275 280 285	
80	cct ggc agg ctt aac gta gat ctg tac gac tat gcc gaa gag atg cca	912
81	Pro Gly Arg Leu Asn Val Asp Leu Tyr Asp Tyr Ala Glu Glu Met Pro	
82	290 295 300	
83	gag atc aag ata aag agt ctc gag gag gtc gca gag tat cta ggc gtg	960
84	Glu Ile Lys Ile Lys Ser Leu Glu Glu Val Ala Glu Tyr Leu Gly Val	
85	305 310 315 320	
86	atg aag aag agt gaa cgc gtt atc atc aat tgg tgg gag att cca gac	1008
87	Met Lys Lys Ser Glu Arg Val Ile Ile Asn Trp Trp Glu Ile Pro Asp	
88	325 330 335	
89	tat tgg gac gac ccg aag aag aga cca cta tta ctg caa tac gcg cgc	1056
90	Tyr Trp Asp Asp Pro Lys Lys Arg Pro Leu Leu Leu Gln Tyr Ala Arg	
91	340 345 350	
92	gac gat gtc cgc gct act tac ggc tta gcc gag aag ata ttg ccg ttt	1104
93	Asp Asp Val Arg Ala Thr Tyr Gly Leu Ala Glu Lys Ile Leu Pro Phe	
94	355 360 365	
95	gct atc cag ttg tcg tac gta aca ggt ctc cca cta gac cag gta ggt	1152
96	Ala Ile Gln Leu Ser Tyr Val Thr Gly Leu Pro Leu Asp Gln Val Gly	

## RAW SEQUENCE LISTING

DATE: 01/30/2002

PATENT APPLICATION: US/10/034,621

TIME: 18:26:31

Input Set : N:\Crf3\RULE60\10034621.raw

Output Set: N:\CRF3\01302002\J034621.raw

97	370	375	380	
98	gcg atg agt gtt ggc ttt cga ctt gaa tgg tac ctg ata cgc gcg gcg	1200		
99	Ala Met Ser Val Gly Phe Arg Leu Glu Trp Tyr Leu Ile Arg Ala Ala			
100	385	390	395	400
101	ttt aag atg aaa gag ctt gtg ccg aac cgc gtt gag cgc cca gaa gag	1248		
102	Phe Lys Met Lys Glu Leu Val Pro Asn Arg Val Glu Arg Pro Glu Glu			
103	405	410	415	
104	act tac cgt ggc gct ata gtt ctt gag ccg ttg aga ggc gtg cac gag	1296		
105	Thr Tyr Arg Gly Ala Ile Val Leu Glu Pro Leu Arg Gly Val His Glu			
106	420	425	430	
107	aat ata gcc gta ctc gac ttt agc tcg atg tac cca aac atc atg ata	1344		
108	Asn Ile Ala Val Leu Asp Phe Ser Ser Met Tyr Pro Asn Ile Met Ile			
109	435	440	445	
110	aag tac aat gtt ggt cct gac acg ctt gtg agg cct ggt gaa aag tgt	1392		
111	Lys Tyr Asn Val Gly Pro Asp Thr Leu Val Arg Pro Gly Glu Lys Cys			
112	450	455	460	
113	ggc gag tgt ggt tgc tgg gag gcc ccg gag gtc aag cac agg ttc cgt	1440		
114	Gly Glu Cys Gly Cys Trp Glu Ala Pro Glu Val Lys His Arg Phe Arg			
115	465	470	475	480
116	agg tgt ccg ccc ggc ttc ttc aag aca gtt ctt gag agg ctg tta gag	1488		
117	Arg Cys Pro Pro Gly Phe Phe Lys Thr Val Leu Glu Arg Leu Leu Glu			
118	485	490	495	
119	ctt cgt aag cgt gtg cgt gct gaa atg aag aag tat cct ccg gat agc	1536		
120	Leu Arg Lys Arg Val Arg Ala Glu Met Lys Lys Tyr Pro Pro Asp Ser			
121	500	505	510	
122	cca gaa tat cga ctg ttg gat gaa agg cag aag gcg ttg aag gtt ctt	1584		
123	Pro Glu Tyr Arg Leu Leu Asp Glu Arg Gln Lys Ala Leu Lys Val Leu			
124	515	520	525	
125	gca aac gct agt tac ggc tac atg ggt tgg agc ggc gct agg tgg tat	1632		
126	Ala Asn Ala Ser Tyr Gly Tyr Met Gly Trp Ser Gly Ala Arg Trp Tyr			
127	530	535	540	
128	tgc agg gag tgc gca aag gct gtc acg gct tgg ggt agg cac ctc ata	1680		
129	Cys Arg Glu Cys Ala Lys Ala Val Thr Ala Trp Gly Arg His Leu Ile			
130	545	550	555	560
131	cgc acc gcc atc aac ata gct cgt aaa cta ggc ctc aag gtg atc tac	1728		
132	Arg Thr Ala Ile Asn Ile Ala Arg Lys Leu Gly Leu Lys Val Ile Tyr			
133	565	570	575	
134	ggt gac aca gat tcg ctc ttc gtg acc tat gat ccg gag aag gtg gaa	1776		
135	Gly Asp Thr Asp Ser Leu Phe Val Thr Tyr Asp Pro Glu Lys Val Glu			
136	580	585	590	
137	aat ttc atc aaa att ata aag gag gag ctg ggg ttc gaa atc aag cta	1824		
138	Asn Phe Ile Lys Ile Ile Lys Glu Glu Leu Gly Phe Glu Ile Lys Leu			
139	595	600	605	
140	gag aag gtg tac aaa cgc tta ttc ttt aca gag gct aag aag agg tac	1872		
141	Glu Lys Val Tyr Lys Arg Leu Phe Phe Thr Glu Ala Lys Lys Arg Tyr			
142	610	615	620	
143	gct ggc ctt ctc gag gac gga cgt ata gat att gtc ggt ttc gag gct	1920		
144	Ala Gly Leu Leu Glu Asp Gly Arg Ile Asp Ile Val Gly Phe Glu Ala			
145	625	630	635	640

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/034,621

DATE: 01/30/2002  
TIME: 18:26:31

Input Set : N:\Crf3\RULE60\10034621.raw  
Output Set: N:\CRF3\01302002\J034621.raw

146	gta cgt ggc gat tgg tgt gaa ctc gcc aag gag gtt cag act aag gtt	1968
147	Val Arg Gly Asp Trp Cys Glu Leu Ala Lys Glu Val Gln Thr Lys Val	
148	645 650 655	
149	gtc gaa ata gta ttg aag acg agt gag gtg aac aag gct gta gag tac	2016
150	Val Glu Ile Val Leu Lys Thr Ser Glu Val Asn Lys Ala Val Glu Tyr	
151	660 665 670	
152	gtc agg aag att gtg aaa gag ttg gag gag ggc aag gtt ccc ata gag	2064
153	Val Arg Lys Ile Val Lys Glu Leu Glu Glu Gly Lys Val Pro Ile Glu	
154	675 680 685	
155	aag ctt gta atc tgg aag acc ctt agt aag cgt ctt gag gag tac aca	2112
156	Lys Leu Val Ile Trp Lys Thr Leu Ser Lys Arg Leu Glu Glu Tyr Thr	
157	690 695 700	
158	acg gag gca cca cac gtc gtt gca gcg aag agg atg ctg tca gca ggc	2160
159	Thr Glu Ala Pro His Val Val Ala Ala Lys Arg Met Leu Ser Ala Gly	
160	705 710 715 720	
161	tac cgg gta agc cca ggc gac aag ata ggg tat gta ata gtg aag ggt	2208
162	Tyr Arg Val Ser Pro Gly Asp Lys Ile Gly Tyr Val Ile Val Lys Gly	
163	725 730 735	
164	ggg ggc cgt atc agt caa aga gca tgg cca tac ttc atg gtc aag gat	2256
165	Gly Gly Arg Ile Ser Gln Arg Ala Trp Pro Tyr Phe Met Val Lys Asp	
166	740 745 750	
167	cct agc cag ata gac gtg acc tac tat gtt gac cac caa atc atc ccg	2304
168	Pro Ser Gln Ile Asp Val Thr Tyr Tyr Val Asp His Gln Ile Ile Pro	
169	755 760 765	
170	gct gca ttg aga ata ctg ggc tac ttt ggc atc acc gag aag aag ctg	2352
171	Ala Ala Leu Arg Ile Leu Gly Tyr Phe Gly Ile Thr Glu Lys Lys Leu	
172	770 775 780	
173	aaa gca agt gca act ggg cag aag act ctc ttc gac ttt cta gcc aag	2400
174	Lys Ala Ser Ala Thr Gly Gln Lys Thr Leu Phe Asp Phe Leu Ala Lys	
175	785 790 795 800	
176	aag agc aag taa	2412
177	Lys Ser Lys	
179	<210> SEQ ID NO: 2	
180	<211> LENGTH: 803	
181	<212> TYPE: PRT	
182	<213> ORGANISM: Pyrolobus fumarius	
183	<400> SEQUENCE: 2	
184	Met Thr Glu Val Val Phe Thr Val Leu Asp Ser Ser Tyr Glu Val Val	
185	1 5 10 15	
186	Gly Lys Glu Pro Gln Val Ile Ile Trp Gly Ile Ala Glu Asn Gly Glu	
187	20 25 30	
188	Arg Val Val Leu Ile Asp Arg Ser Phe Arg Pro Tyr Phe Tyr Ala Leu	
189	35 40 45	
190	Leu Ala Pro Gly Ala Asp Pro Lys Gln Val Ala Gln Arg Ile Arg Ala	
191	50 55 60	
192	Leu Ser Arg Pro Lys Ser Pro Ile Ile Gly Val Glu Asp Asp Lys Arg	
193	65 70 75 80	
194	Lys Tyr Phe Gly Arg Pro Arg Arg Val Leu Arg Ile Arg Thr Val Leu	
195	85 90 95	

## RAW SEQUENCE LISTING

DATE: 01/30/2002

PATENT APPLICATION: US/10/034,621

TIME: 18:26:32

Input Set : N:\Crif3\RULE60\10034621.raw

Output Set: N:\CRF3\01302002\J034621.raw

```

196   Pro Glu Ala Val Arg Glu Tyr Arg Glu Leu Val Lys Asn Val Asp Gly
197           100                      105          110
198   Val Glu Asp Val Leu Glu Ala Asp Ile Arg Phe Ala Met Arg Tyr Leu
199           115                      120          125
200   Ile Asp His Asp Leu Phe Pro Phe Thr Trp Tyr Arg Val Glu Ala Glu
201           130                      135          140
202   Pro Leu Glu Asn Lys Met Gly Phe Arg Val Asp Lys Val Tyr Leu Val
203           145                      150          155          160
204   Lys Ser Arg Pro Glu Pro Leu Tyr Gly Glu Ala Leu Ala Pro Thr Lys
205           165                      170          175
206   Leu Pro Asp Leu Arg Ile Leu Ala Phe Asp Ile Glu Val Tyr Ser Lys
207           180                      185          190
208   Gln Gly Ser Pro Arg Pro Glu Arg Asp Pro Val Ile Val Ile Ala Val
209           195                      200          205
210   Lys Thr Asp Asp Gly Asp Glu Val Leu Phe Ile Ala Glu Gly Lys Asp
211           210                      215          220
212   Asp Arg Lys Pro Ile Arg Glu Phe Val Glu Tyr Val Lys Arg Tyr Asp
213           225                      230          235          240
214   Pro Asp Ile Ile Val Gly Tyr Asn Asn Asn His Phe Asp Trp Pro Tyr
215           245                      250          255
216   Leu Leu Arg Arg Ala Arg Ile Leu Gly Ile Lys Leu Asp Val Thr Arg
217           260                      265          270
218   Arg Val Gly Ala Glu Pro Thr Thr Ser Val His Gly His Val Ser Val
219           275                      280          285
220   Pro Gly Arg Leu Asn Val Asp Leu Tyr Asp Tyr Ala Glu Glu Met Pro
221           290                      295          300
222   Glu Ile Lys Ile Lys Ser Leu Glu Glu Val Ala Glu Tyr Leu Gly Val
223           305                      310          315          320
224   Met Lys Lys Ser Glu Arg Val Ile Ile Asn Trp Trp Glu Ile Pro Asp
225           325                      330          335
226   Tyr Trp Asp Asp Pro Lys Lys Arg Pro Leu Leu Leu Gln Tyr Ala Arg
227           340                      345          350
228   Asp Asp Val Arg Ala Thr Tyr Gly Leu Ala Glu Lys Ile Leu Pro Phe
229           355                      360          365
230   Ala Ile Gln Leu Ser Tyr Val Thr Gly Leu Pro Leu Asp Gln Val Gly
231           370                      375          380
232   Ala Met Ser Val Gly Phe Arg Leu Glu Trp Tyr Leu Ile Arg Ala Ala
233           385                      390          395          400
234   Phe Lys Met Lys Glu Leu Val Pro Asn Arg Val Glu Arg Pro Glu Glu
235           405                      410          415
236   Thr Tyr Arg Gly Ala Ile Val Leu Glu Pro Leu Arg Gly Val His Glu
237           420                      425          430
238   Asn Ile Ala Val Leu Asp Phe Ser Ser Met Tyr Pro Asn Ile Met Ile
239           435                      440          445
240   Lys Tyr Asn Val Gly Pro Asp Thr Leu Val Arg Pro Gly Glu Lys Cys
241           450                      455          460
242   Gly Glu Cys Gly Cys Trp Glu Ala Pro Glu Val Lys His Arg Phe Arg
243           465                      470          475          480
244   Arg Cys Pro Pro Gly Phe Phe Lys Thr Val Leu Glu Arg Leu Leu Glu

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/034,621

DATE: 01/30/2002

TIME: 18:26:33

Input Set : N:\Crf3\RULE60\10034621.raw

Output Set: N:\CRF3\01302002\J034621.raw

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date